Analysis of demographic and socio-economic factors as drivers of deforestation in the Ribeira do Iguape river basin, Brazilian Atlantic Forest: a GIS integration of census and remote sensing data

ABSTRACT:

In this paper, we develop an analysis on the socio-demographic factors associated with land cover change and deforestation processes in the Ribeira Valley watershed, a region that concentrates the largest remnants of the Brazilian Atlantic Forest.
The main objective of the paper is to analyze the role of the different factors (demographic and socio-economic factors, topography, road infra-structure, conservation units) upon land cover change and deforestation processes in the Ribeira Valley region.
In our study, we found out that regression analysis was not appropriate to deal with the independent variables because of the occurrence of strong correlations among them. We decided to work, instead, with a qualitative model that could represent the network of relationships between the independent variables and the deforestation rates (the dependent variable). Concerning this issue, we make a discussion about the limitations of deforestation regression models to address the role of demographic, socioeconomic and infrastructural variables to explain deforestation processes.
The methodology for the analysis integrates socio-demographic data (from the Brazilian demographic census – 1991 and 2000) and land cover change data (from satellite images – Landsat TM), at the level of census tracts, in a geographic information system (GIS).
In sum, our analysis of the factors associated with deforestation in the Ribeira Valley incorporates three important aspects: 1) it presents a wide geographic coverage; 2) it uses very disaggregated spatial unit of analysis (census tracts); and 3) it integrates a large and diverse number of variables for the analysis (census variables, remote sensing/land cover variables and other spatial variables as topography and road infrastructure).
Therefore, maybe the most significant contribution of this study is the application of a methodology that integrates census and remote sensing variables, all information aggregated at the level of census tracts, for the development of an analysis of the associations between socio-demographic factors and deforestation. Hence, this is one of the first studies in the field of "Population and Environment" to use this kind of methodology at the level of census tracts.

KEY WORDS: deforestation; Brazilian Atlantic Forest; geographic information systems; remote sensing; regression models of deforestation.

INTRODUCTION

In this paper, we present an analysis of the drivers of deforestation in the Brazilian Atlantic Forest remnants, specifically in the Ribeira de Iguape River Basin, which covers an area of approximately 28 thousand square kilometers, 62% of it located in São Paulo State (involving 23 municipalities) and 38% in Parana State (distributed in 9 municipalities). Our study will focus on the São Paulo portion of the Ribeira de Iguape River Basin.

Recognized for its biological diversity, the so-called Ribeira Valley region covers coastal, floodplain and mountain areas and represents the most significant fragment of the Atlantic Forest in Brazil. Twelve federal and state conservation units are located in this area. The region has been considered one of the most important international conservation priorities by agencies such as the International Union for the Conservation of Nature (IUCN), and the UNESCO/MAB Program.

Given these elements, the main objective of this paper is to analyze the relationships (associations, correlations and causalities) between land cover change processes (recent deforestation of Atlantic Forest remnants) and demographic and socio-economic factors, topography, road infra-structure and conservation units, at the level of census tracts.

In our study, we found out that regression analysis was not appropriate to deal with the independent variables because of the occurrence of strong correlations among them. Consequently, due to these restrictions to work with several independent variables in regression models, we decided to work, instead, with a qualitative model that could represent the network of relationships between the independent variables and the deforestation rates (the dependent variable). The major advantage of adopting a qualitative model is the possibility to map and represent graphically the diversity of factors associated with deforestation in the Ribeira Valley. Concerning this issue, we make a discussion about the limitations of deforestation regression models to address the role of demographic, socioeconomic and infrastructural variables to explain deforestation processes.

In order to build a qualitative model and achieve our objective, we developed a methodology that integrates socio-demographic data (from the Brazilian demographic census – 1991 and 2000) with land cover change data (from satellite images – Landsat TM, years 1990 and 1999) at the level of census tracts, within a geographic information system (GIS).

In sum, our analysis of the factors associated with deforestation in the Ribeira Valley incorporates three important aspects: 1) it presents a wide geographic coverage; 2) it uses very disaggregated spatial unit of analysis (census tracts); and 3) it integrates a large and diverse number of variables for the analysis (census variables, remote sensing/land cover variables and other spatial variables as topography and road infrastructure).

In this sense, maybe the most significant contribution of this study is the application of a methodology that integrates census and remote sensing variables, all information aggregated at the level of census tracts, for the development of an analysis of the associations between socio-demographic factors and deforestation. Hence, this is one of the first studies in the field of “Population and Environment” to use this kind of methodology at the level of census tracts.

This paper is divided into five sections, described as follows.

In the first session, we make a brief review and discussion of the literature on drivers of deforestation, focusing on the limitations of regression models to address the role of demographic, socioeconomic and infrastructural variables to explain deforestation processes.

The second section presents a general description and characterization of the Ribeira Valley region. In the third section, we explain the methodology and databases used for the analysis, which involves, mainly, the integration of census and remote sensing data at the level of census tracts within a Geographic Information System (GIS).

In the forth section, we carry out an analysis of the factors (drivers) associated with deforestation processes in the Ribeira Valley. We analyze the relationships (associations and correlations) between deforestation and the factors mentioned above (demographic and socio-economic factors, topography, road infra-structure and conservation units), at the level of census tracts. At the end of this section, we propose a qualitative model of deforestation in the Ribeira Valley, representing the network of relationships between the independent variables and the deforestation rates (the dependent variable).

Finally, in the fifth section, we make a discussion of the results and present final remarks.

BRIEF REVIEW AND DISCUSSION OF THE LITERATURE ON DRIVERS OF DEFORESTATION: THE LIMITS OF REGRESSION MODELS

In this first session, we make a brief review and discussion of the literature on drivers of deforestation, focusing on the limitations of regression models to address the role of demographic, socioeconomic and infrastructural variables to explain deforestation processes. Our discussion is based mostly on two important reviews of models and case studies of tropical deforestation, presented in Kaimowitz & Angelsen (1998) and Geist & Lambin (2001).

One of the key contentious issues in global environmental change research relates to the major human causes and drivers of land cover change in different geographical and historical contexts. In the same way, there has been considerable improvement concerning inter-disciplinary research on the so-called human dimensions of environmental change, integrating methodologies, databases and research groups from the natural and social sciences (Liverman et al., 1998). The land use and land cover change processes have been one of the main topics in this area working with an integrated approach for research, incorporating demographic and socioeconomic analysis to the land use cover change research agenda (IGBP-IHDP LUCC Project).

The literature on models and case studies of tropical deforestation reflects the great effort of research and modeling in order to identify the factors associated with deforestation and to explain its causative patterns. In this sense, a great effort has been made to identify and explain the causes and drivers of deforestation, especially in tropical regions (Allen & Barnes, 1985; Walker, 1987; Rudel, 1989; Lambin, 1994, 1997; Lambin et al., 2001; Sponsel *et al.*, 1996; Rudel & Roper, 1996, 1997; Kaimowitz & Angelsen, 1998; Mather & Needle, 2000; Geist & Lambin, 2001, 2002).

Broadly speaking, the studies about land cover change and deforestation can be classified in two large groups: 1) the case studies and 2) the formal models. There is a striking contrast between the complexity of descriptions of land cover change and deforestation processes for specific case studies and the relative simplicity of the mechanisms represented in formal models. However, in spite of the great precision and detail in describing deforestation processes, one of the case studies´ major drawback is the limited geographical coverage and the unfeasibility of generalization for larger areas, such as a region or a river basin (Lambin, 1997).

On the other hand, formal models usually work with regional, national or even global scales. Additionally, most of these models are empirical and one of the methodologies most used are regression analysis. Yet, deforestation regression models have strong limitations to address the role of demographic, socioeconomic and infrastructural variables to explain the deforestation processes that they are studying. One of the basic hypotheses in linear regression models is that independent variables are not correlated. In land use/cover data, this hypothesis is usually not true. Moreover, land use/cover data have the tendency to be spatially autocorrelated, as land-use changes in one area tend to propagate to neighboring regions (Overmars, Koning, & Veldkamp, 2003; Aguiar, Câmara &Escada, 2007).

For Scrieciu (2007: 610), "regression analyses that seek to explain deforestation at the global level need to be carefully scrutinized and checked to ensure that they meet standard statistical tests". In this sense, "tropical deforestation might ultimately depend upon case-specific factors, and further research may benefit to a greater extent and may render more effective policy suggestions if conducted at a more disaggregated level".

In sum, one of the most important limitations on deforestation regression models is the occurrence of strong correlations among the independent variables. As a consequence, the results found by deforestation regression models are problematical and contradictory, as one can see in the brief review of models of deforestation discussed below.

In the following, the discussion is organized according to the groups of factors (drivers) associated with deforestation, as shown in the qualitative model that we built for the Ribeira Valley (see the fourth session). All the factors we have analyzed for the Ribeira Valley are pointed out as drivers of tropical deforestation in the models and case studies described in the two reviews of the literature.

Demographic factors (population size, density and growth)

Much of the land use/cover change literature accepts that population change and distribution is a significant driver of global deforestation. "For example, Mather et al. (1998) estimate that population explains approximately half of the variation in deforestation worldwide while Allen and Barnes (1985) consider it the primary cause of the planet's deforestation" (Carr, 2004: 587).

In this sense, the majority of the global regression models of deforestation show that demographic factors (mainly population size, density and growth) are the most important drivers of tropical deforestation (Mather & Needle, 2000; Allen & Barnes, 1985). In fact, "much of the work attempting to quantify causal linkages between population and deforestation has been limited by an early focus on global-level regression analyses" (Carr, Suter & Barbieri, 2005: 93). Nevertheless, these regression models tend to suffer from data limitations, particularly in terms of which variables exist in extant data bases, and how well they measure what they purport to measure (Geist& Lambin 2002).

Besides, the conclusions presented in the two important reviews of models and case studies of deforestation show that population, specially its growth, is not the main factor associated with deforestation in the regional and local levels (Kaimowitz & Angelsen, 1998; Geist & Lambin, 2001). In the review of 152 case studies, at regional and local levels, carried out by Geist & Lambin (2001), population factors are reported as drivers of deforestation in 93 cases (61% of all) but, although they present a significant impact on deforestation, they are not as important as other factors like socioeconomic, political, technological, institutional and socio-cultural factors.

Among the case studies reviewed by Geist & Lambin (2001), population is never seen as a single factor associated with deforestation, i.e., its effect on deforestation is always derived from inter-linkages with other drivers (socioeconomic, political, technological etc.). For Carr, Suter & Barbieri, (2005: 91), "the role of population in driving deforestation is complex, and characterized by a host of spatial and temporal contingencies. Population effects are nearly always mediated through political, economic, and ecological factors".

The same pattern is found in our model for the Ribeira Valley, with many factors affecting deforestation along with population, such as road network, proximity to urban centers and socioeconomic conditions. In particular, we have seen that the population density in rural census tracts is highly correlated with road network density and proximity to urban centers. In many deforestation models reviewed by Kaimowitz & Angelsen (1998), the population density, at regional and local levels, has a strong correlation with other factors, especially with road infrastructure and access to urban markets. Therefore, the strong correlation found between population density and deforestation in the rural census tracts of the Ribeira Valley could be seen as consequence of the effect of roads and urban centers on deforestation.

Another aspect of confluence between our results for the Ribeira Valley and the two mentioned reviews is the smaller importance of population growth upon deforestation processes. In our analysis for the Ribeira Valley, the correlation between population growth and deforestation is the weakest among all other factors associated with deforestation at the level of census tracts.

In fact, the evidences found in the models about the association between population growth and deforestation are very weak. The models and case studies reviewed do not support the idea – as put forward, for example, by Allen & Barnes (1985) – that population growth is the primary cause of deforestation, particularly population growth derived from high fertility levels (Geist & Lambin, 2001, 2002; Lambin et al., 2001).

Income levels and socioeconomic conditions

Most global deforestation regression models reviewed by Kaimowitz & Angelsen (1998) find a positive association between higher national per capita income and greater levels of deforestation (Capistrano & Kiker, 1995; Rock 1996). For Carr, Suter & Barbieri, (2005: 93), "cross-national statistical analyses popular in the 1980s and early 1990s managed to correlate national level rates of deforestation with economic and population growth rates (e.g. Rudel & Roper, 1997)". Though these studies demonstrated these associations with deforestation at aggregate scales, further studies have shown that at local to regional scales there may be no such association (Rindfuss, Turner, Entwisle, & Walsh, 2004).

Additionally, deforestation models based on a regional level and that have attempted to measure the relationship between income levels and deforestation have obtained contradictory results. On one hand, higher income levels are expected to increase pressure on forests by rising the demand for agricultural and forest products and by stimulating new access to virgin forests. But, on the other hand, in regions with higher per capita income (and therefore with higher wages), the logging and agriculture activities related to deforestation become less profitable and hence lead to lower forest depletion (Kaimowitz & Angelsen, 1998). Moreover, deforestation may contribute to increase the population income levels, which could imply a causal relationship working in the opposite direction.

In our qualitative model for the Ribeira Valley, income levels and other socioeconomic conditions (e. g. schooling and sanitation) present positive associations with deforestation. However, it was difficult for us to establish a causal relationship and, especially, its direction. In some cases, the deforestation could have generated income, therefore improving socioeconomic conditions of the rural population in the census tracts.

Poverty

The review of deforestation models carried out by Kaimowitz & Angelsen (1998) oppose the conventional wisdom that says that rural poverty is an important driver of deforestation in the tropics, showing that "there is little empirical evidence on the link between deforestation and poverty. If forest clearing requires investment, rich people may in fact be in a better position to clear new forest land" (Angelsen & Kaimowitz, 1999: 92).

In the review carried out by Geist & Lambin (2001), poverty appears as a factor associated with deforestation in only 15% of the case studies reviewed, located mainly in Asia.

Our results for the rural census tracts of the Ribeira Valley go in the same direction. Among the census tracts of the Ribeira Valley, the degree of poverty of the heads of households is *negatively* associated with deforestation.

Road infrastructure and access to urban markets

Deforestation models reviewed by Kaimowitz & Angelsen (1998) find that greater access to forests and markets accelerates deforestation. "Spatial regression models are well suited for studying the effects of access. Models of this type for Belize (Chomitz & Gray, 1996), Cameroon (Mertens & Lambim, 1997) and Costa Rica (Rosero-Bixby & Palloni, 1998) all show a strong relation between roads and deforestation. These results are also supported by nonspatial regression models from Brazil (Pfaff, 1997) and Ecuador (Southgate et al., 1991). Most studies show that forest clearing declines rapidly beyond distances of 2 or 3 kilometers from a road" (Angelsen & Kaimowitz, 1999: 85).

Concerning access to urban markets, Chomitz & Gray (1996), in a case study about Belize, show that areas closer to urban markets have less forest cover. Mertens & Lambim (1997), in a case study about Cameroon, show that deforestation rates fall remarkably beyond a 10 km distance from an urban center.

In the review of 152 case studies carried out by Geist & Lambin (2001), the presence of roads, especially road construction, is an important proximate cause of tropical deforestation, appearing in 61% of all case studies reviewed[1].

In our qualitative model for the Ribeira Valley, the proximity to urban centers is the factor (independent variable) presenting the second highest positive correlation with deforestation and is also highly correlated with population density. The road network density also shows a positive correlation with deforestation but the association is not as strong as with other independent variables. However, there is a very high negative

[1] Geist & Lambin (2001) compute any type of road, including unfinished penetration or feeder roads, logging roads and oil and mining company roads.

association between road network density and the percentage of forest remnants in the census tract, suggesting that there was an important effect of road construction on deforestation in the past decades.

Topography

Topography is mentioned in only 5% of the case studies on land cover change and deforestation processes reviewed by Geist & Lambin (2001), most of them in Latin America. Those studies show that flat and gently sloping areas favor deforestation. However, the topography plays a very important role in the preservation of Atlantic Forest remnants in the Ribeira Valley. The topography (elevation range inside the census tract) has a strong negative association with deforestation.

Conservation Units

At last, it is interesting to highlight an important study that analyzed 93 protected areas in 22 tropical countries. This study concluded that the majority of the conservation units, especially the ones with stronger land use restrictions as parks and reserves, have been very successful in the protection of tropical forests, a fairly surprising finding in the context of chronicle lack of economic resources for environmental protection and the great land cover pressure in these areas. The conservation units have been particularly efficient in avoiding deforestation processes, which is considered the major threat to biodiversity in the tropics (Bruner et al., 2001).

Our analysis for the Ribeira Valley also shows that the conservation units play a very important role in the preservation of Atlantic Forest remnants. The census tracts inside conservation units show significantly lower deforestation rates and significantly higher proportion of forest remnants than the census tracts outside conservation units.

Curiously, only two models reviewed by Kaimowitz & Angelsen (1998) mention conservation units or protected areas as factors associated with deforestation; these two models conclude that the protection status reduces the probability of an area being deforested.

THE RIBEIRA VALLEY REGION

The Ribeira de Iguape River Basin covers an area of approximately 28 thousand square kilometers, 62% of which is located in São Paulo State and 38% in Parana State. The Ribeira de Iguape River extends through 470 kilometers and is the last river in São

Paulo State that has not been changed by dams, hydroelectric powders or other large engineering interventions. The Ribeira Valley region covers coastal, floodplain and mountain areas, and represents the most significant fragment of Atlantic Forest remaining in Brazil with twelve federal and state conservation units. The region has been considered one of the most important international conservation priorities by agencies such as the International Union for the Conservation of Nature (IUCN), and the UNESCO/MAB Program. Additionally, a single Landsat TM scene (220/77) covers almost the entire study area.

The Ribeira Valley encompasses the 'Estuarine Lagunar Complex of Iguape and Cananéia' (also known as 'Lagamar'), a vast estuary over 100 kilometers long and one of the five largest reproductive sites for South Atlantic marine species. This estuary is constituted of many coastal islands and extensive mangrove swamps, forming one of the most productive ecosystems in the world, where many species of fish and crustaceans feed and reproduce. The Ribeira de Iguape is the estuary's major river. There is, therefore, an inherent unity between the estuary and the Ribeira de Iguape River Basin.

The Ribeira Valley Region is one of the six areas that form the Atlantic Forest Biosphere Reserve[2]. It is the most preserved area and concentrates the most extensive continuous remnants of tropical forest and associated ecosystems of all the six areas. The International Union for the Conservation of Nature (IUCN) considers this region as one of the highest priority areas for conservation in the world.

The Ribeira Valley region as a whole (including the São Paulo and Parana States areas) presents an extraordinary environmental heritage with over 2.1 million hectares of forests or approximately the equivalent to 21% of all the remaining Atlantic Forest in Brazil. It also includes 150 thousand hectares of *restingas* (beach vegetation) and 17,000 hectares of mangroves - all very well preserved – besides being one of the most important speleological heritages in the country. The region is also rich in ethnical and cultural terms as the population living in the Ribeira Valley includes communities of native Indians, *caiçaras* (community of seaside inhabitants), *quilombolas* (communities formed by descendants of black slaves) and small agricultural families, forming a cultural diversity that is quite rare to be seen in places so close to developed or urbanized regions (ISA, 1998; Lino, 1992).

---

[2] In the context of Unesco's *Man and the Biosphere – MAB – Program.*

In contrast with its rich environmental heritage, the Ribeira Valley is the least developed region of São Paulo, Brazil's most industrialized and wealthiest state. It presents the lowest rate of urbanization, the lowest levels of family income, years of schooling of the population and the highest infant mortality and fertility rates. Its economy is based on agriculture (bananas, tea), mining and extraction of forest products (hearts of palm). Historical factors, difficulties of access and natural conditions adverse to economic activities guaranteed, until recently, a relative isolation and preservation of natural resources in the Valley. This is valid even considering the proximity of the Ribeira Valley to urbanized areas. The regional capital, Registro, for example, is located less than two hundred kilometers from the São Paulo Metropolitan Area (ISA, 1998).

In other important aspects of the demographic dynamics of this region, the contrast with the rest of the State of São Paulo is also impressive. The infant mortality rate of the region in 1997 was 31.68 deaths per 1,000 live births, one-third higher than the average for the State. The region is perhaps the last in the State that has not gone through the "epidemiological transition". Its mortality profile is typical of a pre-industrial era, with a predominance of infectious and contagious over chronic and degenerative diseases. In terms of fertility, the regional health district (known as 'DIR of Registro') reports the highest rate in the State, 2.68 versus 2.26 children per woman in the State average (Hogan et al., 1999).

The Ribeira Valley is also the least urbanized region of the State – 65% vs 90% in the State average. Its population growth rate has been historically low, a situation reinforced nowadays by out-migration. Analysis of migration trends reveals three important aspects. First, the Ribeira Valley's net out-migration is considerably high, indicating a lack of perspectives for jobs and socioeconomic insertion. Second, a large part of this mobility is restricted to surrounding municipalities. Adding these movements to the Valley's significant *internal* mobility, we see that the precariousness of employment provokes a considerable volume of population circulation. Third, the region's most 'qualified' residents (or the least unqualified) end up migrating to other areas and, consequently, the least qualified remain settled. In many of the municipalities of the Ribeira Valley, educational levels are extremely low (Hogan et al., 1999).

The Ribeira Valley, on the other hand, has had one of the most intensive political-institutional efforts to create protected areas for the last 30 years in Brazil. The distinct restrictions on land use (including forest use) imposed by the creation of conservation units, the lack of economic alternatives and infrastructure, together with land ownership conflicts have motivated diverse scenarios of environmental changes, making this region an ideal socioenvironmental microcosm for the study of processes of deforestation and land use and land cover change, derived from demographic and socio-economic dynamics and institutional (environmental) restrictions (Brondizio, 2000; ISA, 1998; Hogan et al., 1999).

METHODOLOGIES FOR THE INTEGRATION OF CENSUS AND REMOTE SENSING DATA: A STUDY ON THE RIBEIRA VALLEY

As mentioned in the introduction, the general methodology for the analysis of the factors (drivers) of deforestation in the Ribeira Valley is the integration of socio-demographic data (from the Brazilian demographic census – 1991 and 2000) and land cover change data (from satellite images – Landsat TM, years 1990 and 1999), at the level of census tracts, in a geographic information system (GIS).

The main methods used to integrate census and remote sensing data are:

i. classification of two Landsat TM images of the Ribeira Valley (scene 220/77, years 1990 and 1999);

ii. construction and organization of a demographic and socioeconomic database, from Brazilian census of 1991 and 2000, at the level of municipalities and census tracts;

iii. creation and organization of digital layers of road network, urban centers, topography, rivers, conservation units and municipal and census tracts boundaries;

iv. development of a GIS structure, integrating the three main sources of data mentioned above (satellite images, census data and digital layers);

v. creation of variables for land cover, topography, conservation units and road infrastructure, within a GIS.

Next we make a brief description of each of the methods listed above.

i. Classification of satellite images and creation of a transition matrix

We classified two Landsat TM images of the Ribeira Valley (scene 220/77, years 1990 and 1999). For each of the two images, we have distinguished seven land cover classes, which are: water, forest, mangrove, planted forest and non-forested areas (which include crops, pasture, bare soil and urbanized areas). However, for the purpose of this paper, we looked only at three main land cover types: water, non-forest and forest (which includes primary and secondary forest, mangrove and planted forest).

After the classification was accomplished, we made a change detection analysis, in order to capture and quantify land cover changes between the two dates of the images (time period of 1990-1999). The most important land cover trajectories (or changes) we looked at were deforestation and preservation of forest remnants.

It is important to say that there is a great temporal correspondence between the land cover change/deforestation time period (1990 to 1999) and the census variables (1991 and 2000). This correspondence allows a fine temporal association between land cover change and socio-demographic dynamics.

For more details about the procedures for the classification of the satellite images and the generation of the land cover change variables, see **Appendix 1**.

ii. Demographic and socioeconomic database

We organized a database with demographic and socioeconomic variables from the Brazilian census of 1991 and 2000, at the level of municipalities and census tracts ('setores censitários').

The census data are geo-referred to digital layers of municipalities and census tracts. This is an essential feature in order to make possible the spatial distribution of census data and the integration with remote sensing data and other spatial variables, all within a GIS structure.

iii. Creation and organization of digital layers

We built and organized biophysical, infrastructure and political-administrative maps (GIS vector and raster layers) with road network, urban centers, topography (Digital Elevation Model - DEM), rivers, conservation units and municipal and census tracts boundaries.

The digital layers, as well as the two Landsat images, are projected in the *Universal Transverse Mercator* (UTM) coordinate system, using the SAD 69 datum. Some of the layers were especially built by us at the ACT – Indiana University and others were given by the Brazilian NGO *Instituto Socioambiental*.

iv. Development of a Geographic Information System (GIS)

We developed a GIS structure integrating three main sources of data:

1) Land cover maps derived from Landsat TM (scene 220/77) images (years 1990 and 1999).

2) Demographic and socioeconomic data from the Brazilian national census (years 1991 and 2000) at municipality and census tracts levels.

3) Biophysical, infrastructure and political-administrative maps (GIS vector and raster layers) with road network, urban centers, topography (Digital Elevation Model - DEM), rivers, conservation units and municipal and census tracts boundaries.

v. Creation of variables for land cover, topography, conservation units and urban and road infra-structure, integrated into a GIS

Besides the census variables, we also created a set of spatial variables integrated into a GIS. These variables were organized in four groups, which are: 1) land cover variables; 2) zoning categories and 'conservation units' variables; 3) topographic variables and 4) urban and road infra-structure variables.

The land cover change variables were created by a two-step procedure. The first step was overlaying the municipalities and census tracts layers with the already classified Landsat images (land cover maps). The second procedure was to extract land cover classes, which were then aggregated to the spatial units of analysis, namely census tracts and municipalities.

This procedure enabled us to estimate the total area (and the percentages) for each land cover class within every municipality and census tract of the Ribeira Valley, therefore, accomplishing the integration of census and remote sensing data. It became possible to estimate, for example, that the municipality of 'Eldorado' has 121 thousand hectares of forest (or 73% of the area of the municipality), 28 thousand ha without forest cover (17% of the municipality) and that the area deforested between 1990 and 1999 was of 10 thousand ha (or 6% of the municipality).

Figure 1 shows how the land cover variables were created by overlaying the municipalities/census tracts layers with the land cover change maps.

The variables regarding zoning categories and conservation units were created by overlaying the conservation units layer with the municipalities/census tracts layers. For the municipalities, we estimated the area (and the percentages of the area) of the municipality under each type of zoning category, i.e., forest reserves/parks, areas of environmental protection (known as 'APAs') and non-protected areas. For the census tracts, we classified each one according to the zoning category that it belonged to: parks, APAs or non-protected areas. This was a very important accomplishment as it became possible to test the role of conservation units on deforestation processes, at the level of census tracts.

Figure 2 shows how the zoning categories/conservation units variables were created by overlaying the conservation units with the municipalities/census tracts layers.

The topographic variables were created with the support of a Digital Elevation Model (DEM). we overlayed the municipalities/census tracts layers with the DEM and calculated a series of descriptive statistics about the topography of each census tracts and municipalities, such as average elevation and slope, minimum, maximum and range of elevation/slope.

It is important to note that the topographic variables are important for our analysis, because the elevation and slope play a very significant role in the processes of land cover change and deforestation in the Ribeira Valley.

Figure 3 shows how the topographic variables were created by overlaying the municipalities/census tracts layers with the Digital Elevation Model.

The last procedure was to create urban and road infra-structure variables by overlaying maps (layers) of urban centers and roads with the municipalities/census tracts layers. First, we created buffers around the main roads and urban centers. For the roads, we tested buffers of 100, 200, 500, 800 and 1,000 meters. For the urban centers, we tested buffers of 1, 3, 5, 8 and 10 kilometers. Afterwards, these buffers were overlayed with the municipalities/census tracts layers and thus we were able to calculate the area (and the percentages of the area) inside the buffers of roads and urban centers for every municipality and census tract. Therefore, these variables were considered as a proxy of access to urban markets and road infra-structure.

Figure 4 shows how the urban and road infra-structure variables were created, through the overlaying of the buffers of urban centers and roads to the municipalities/census tracts layers.

After creating all the spatial variables mentioned above, the database was completed, so for every municipality and census tract the following groups of variables were available:

1) census variables (demographic and socioeconomic variables) (years 1991 and 2000).
2) land cover variables (years 1990 and 1999).
3) zoning categories / conservation units variables.
4) topographic variables.
5) urban and road infra-structure variables.

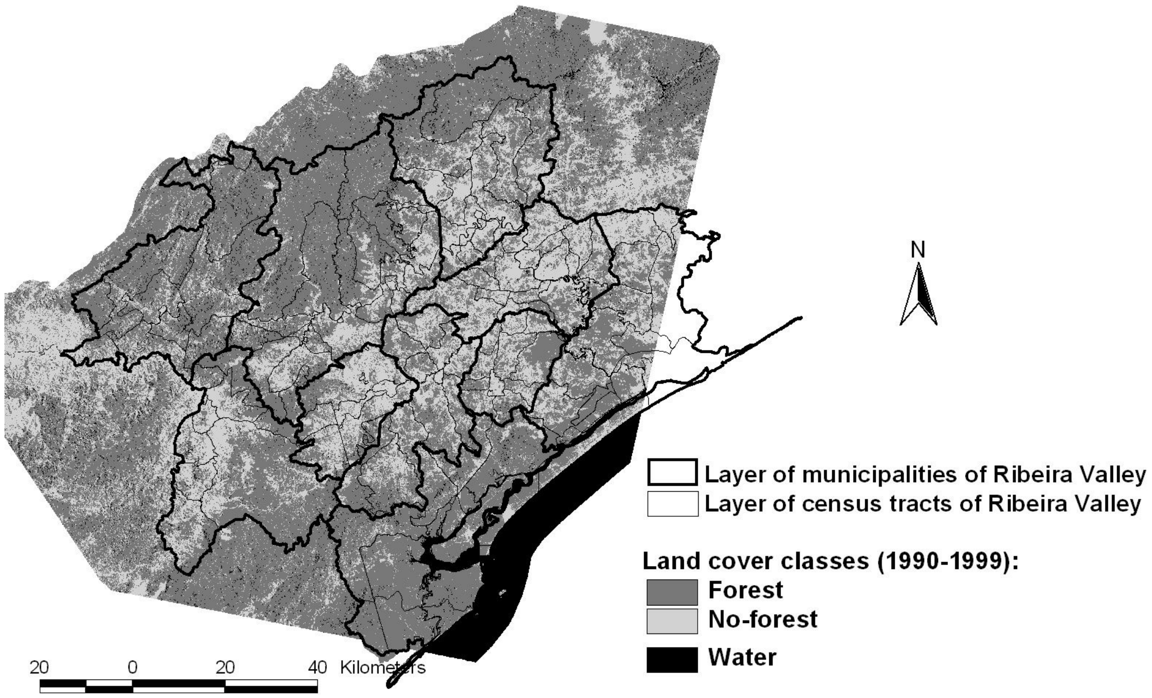

FIGURE 1. Overlaying of the municipalities/census tracts layers of the Ribeira Valley with the land cover maps (classified Landsat TM images)

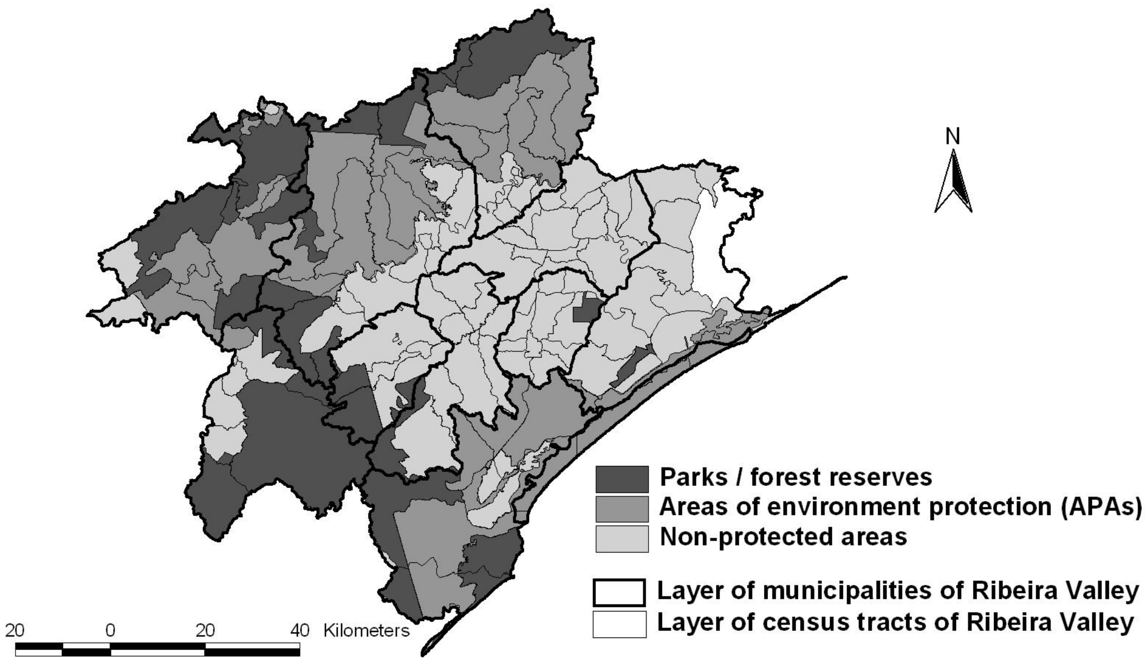

FIGURE 2. Overlaying of the conservation units layer with the municipalities/census tracts layers of the Ribeira Valley

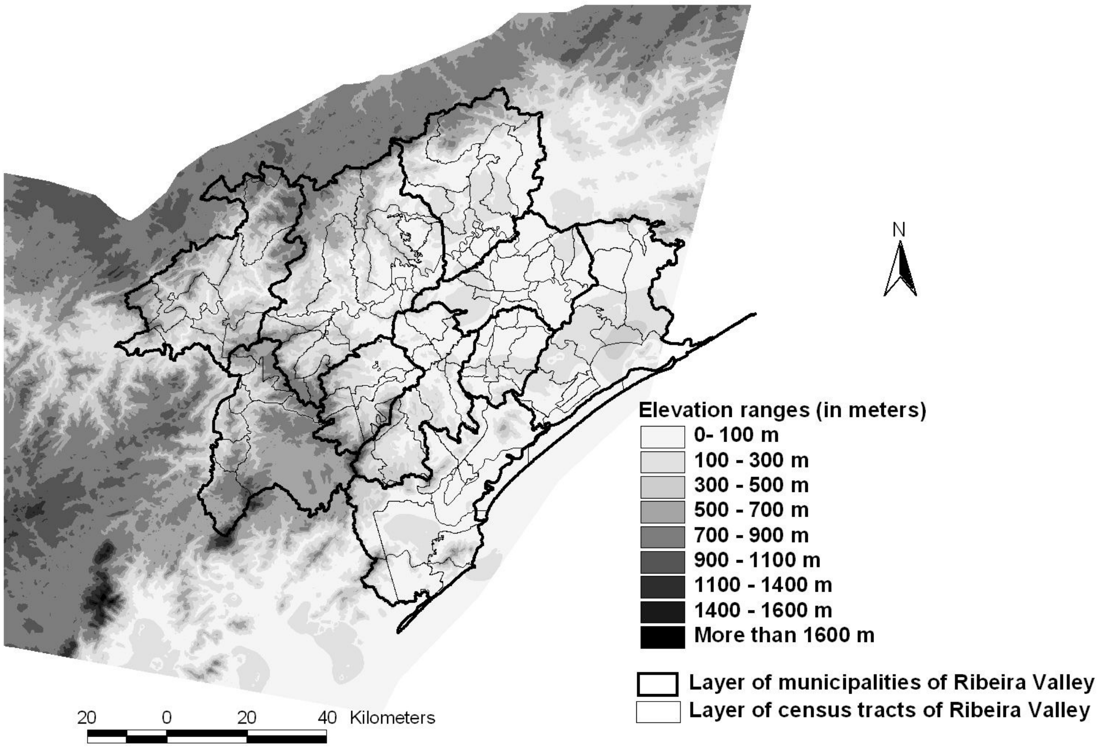

FIGURE 3. Overlaying of the municipalities/census tracts layers with the Digital Elevation Model (DEM) of the Ribeira Valley

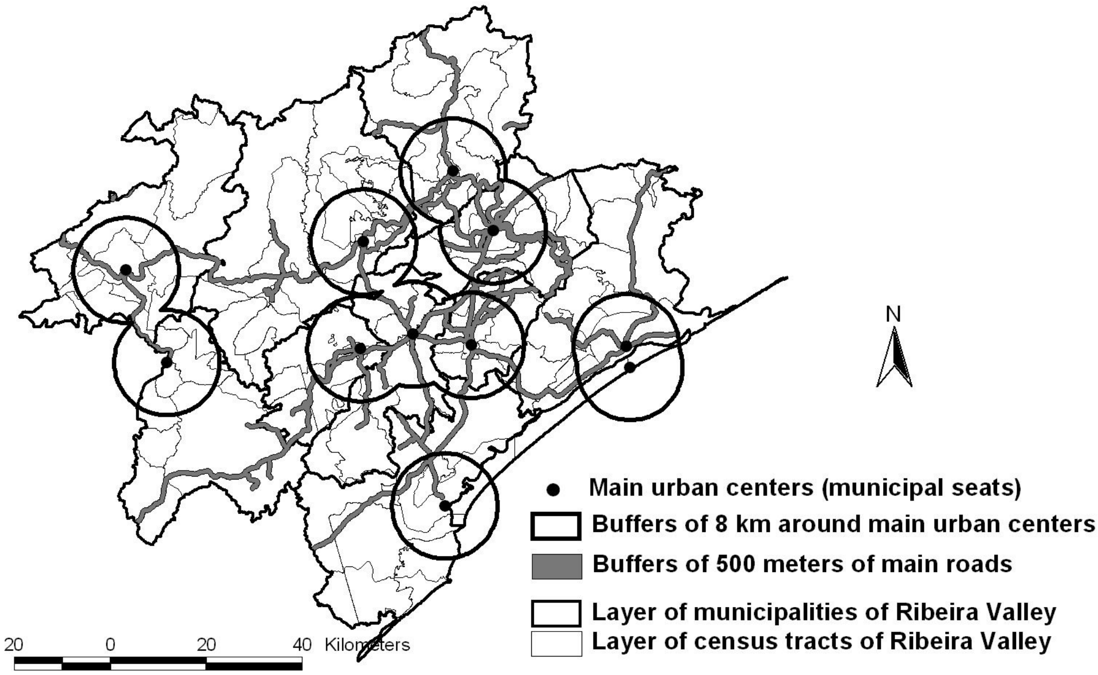

FIGURE 4. Overlaying of the buffers of urban centers and roads to the municipalities/census tracts layers of the Ribeira Valley

ANALYSIS OF THE FACTORS (DRIVERS) ASSOCIATED WITH DEFORESTATION PROCESSES IN THE RIBEIRA VALLEY

In **table 1**, one can see that the factors positively associated with deforestation in the Ribeira Valley are the following:

- demographic factors (population size, density and growth);
- socio-economic conditions (levels of income, schooling and sanitation);
- access to urban markets and road infra-structure (proximity to urban centers and road network).

On the other hand, the factors negatively associated with deforestation are:

- poverty (percentage of head of households earning less than 1 minimum wage);
- topography (range of elevation inside the census tract);
- presence of conservation units.

As shown in table 1, the factor (independent variable) with the strongest positive association with recent deforestation is population density, presenting a linear correlation (Pearson coefficient) of 0.486. The second strongest factor associated with deforestation is the proximity to urban centers, with a linear correlation of 0.452.

Moreover, population density and proximity to urban centers are also highly correlated with each other[3]. So, it is possible to say that this quite strong correlation between population density and deforestation might be a consequence of the effect of the proximity to urban centers on the deforestation rates of the census tracts.

The socio-economic conditions of the population in the census tract (levels of income, schooling and sanitation) have also quite strong positive associations with deforestation, with linear correlations of 0.356 (income), 0.395 (schooling) and 0.405 (sanitation). One possible explanation for these associations is that better socio-economic conditions (therefore higher levels of income and consumption) imply on higher demands for agricultural and forest products and also a greater availability of economic resources to invest in agriculture activities, therefore increasing the probability of deforestation. But the causative pattern can also work in the opposite direction, i.e., the income generated by deforestation could have improved the socio-

[3] The linear correlation between population density and proximity to urban centers is significant and positive, with a value of 0.538.

economic conditions of the population living in census tracts with higher deforestation rates.

TABLE 1.

Linear correlation coefficients between the factors associated with deforestation (selected independent variables) and the rate of deforestation in 1990-1999. Rural Census Tracts of the Ribeira Valley.

Factors associated with deforestation (selected independent variables)	Linear correlation with rate of deforestation (Pearson coefficient)
Population Density in 2000	0.486(**)
Percentage of the census tracts area within 10 Km from the nearest seat of municipality	0.452(**)
Percentage of households with bathroom (in 2000)	0.405(**)
Average years of schooling of the heads of households (in 2000)	0.395(**)
Population size of census tract in 2000	0.362(**)
Percentage of census tracts area within the 1 Km buffers of road network	0.361(**)
Census tracts average income of the head of household (in minimum wages, 2000)	0.356(**)
Population Growth rate (1991-2000)	0.324(**)
Variation of the elevation in the census tract	– 0.490(**)
Percentage of head of households in poverty (in 2000)	– 0.382(**)
Restrictions in land use (presence or not of conservation units)	[0.452] (1)

Source: IBGE. Demographic Census 1991 and 2000 and Landsat TM images, 1990 and 1999 (scene 220/77).

** Significant correlation ($p < 0.01$).

(1) Coefficient of determination (R-square) in the analysis of variance.

After socio-economic conditions, the population size of the census tract is the variable with strongest association with deforestation (linear correlation of 0.362). The road network density is also positively associated with deforestation (Pearson's coefficient 0.361). Besides that, road network has a very high correlation with population density[4], revealing the important role of roads in the spatial distribution of the population, therefore suggesting that population density might be only a proxy of the effect of road network on deforestation in the Ribeira Valley.

[4] The linear correlation between road network and population density of the census tract is significant and positive, with a value of 0.699.

At last, one can see that the population growth rate of the census tract has the lowest positive association with deforestation, among all the factors (independent variables) shown in table 1, with a linear correlation of 0.324. Different from its density, the rate of population growth does not seem to have an important effect on recent deforestation in the Ribeira Valley region.

In table 1, one can also see that the factors negatively associated with deforestation are the degree of poverty[5], the topography[6] and the presence of conservation units.

The degree of poverty presents a negative correlation with deforestation (Pearson coefficient of –0.382), showing that in the census tracts that present a higher percentage of poor head of households, the rates of deforestation are lower.

The topography has a very important negative effect on deforestation, with a negative correlation between deforestation and the range of elevation of the census tract (Pearson coefficient of –0.490). It is important to highlight that, in absolute terms, the range elevation of the census tract is the variable with the strongest correlation with deforestation, even higher than the correlation between population density and deforestation.

The presence of conservation units has also an important negative effect on deforestation rates. In other words, higher rates of deforestation occur in census tracts outside conservation units.

Based on our findings on the relationships between deforestation rates and the selected independent variables (factors) we can say that the rural census tracts with higher rates of deforestation have bigger population size and density, are located close to urban centers (within a 10 km radius), have a more dense road network, better socio-economic conditions and higher rates of population growth. Moreover, the census tracts with higher deforestation rates are located in areas with smoother topography and outside conservation units, and present lower levels of poverty.

In sum, the factors with the strongest positive association with deforestation are the population density and the proximity to urban centers, which are also positively

[5] Percentage of head of households earning less than 1 minimum wage (100 dollars) per month in the census tract.
[6] Range of elevation (minimum and maximum) inside the census tract.

correlated with each other. The factors with the strongest negative association with deforestation are topography (range of elevation of the census tract) and the presence of conservation units.

Besides that, the fact that many factors (independent variables) shown in table 1 are highly correlated with each other indicates that it is not possible to consider each variable as a single or isolated factor associated with deforestation, but instead it should be seen as a "network of relationships", with direct and indirect effects on deforestation processes in the Ribeira Valley.

Based on this "network of relationships" among the independent variables and the deforestation rates, we propose a qualitative (or graphic) model of correlation and causality between socio-demographic factors, topographic and infra-structure attributes, presence of conservation units and recent deforestation in the Ribeira Valley (see **figure 5**).[7]

Figure 5 shows the results for the qualitative model. It is possible to note a web of associations between a set of different factors (demographic, socio-economic, topography, road and urban infra-structure and conservation units) and deforestation processes in the Ribeira Valley.

Despite the fact that this qualitative model is derived from the observed correlations between the independent variables (the factors shown in table 1) and deforestation (the dependent variable), one can still infer some causal relationships from these correlations. Some of the causal associations are very evident, such as the association between topography and conservation units with deforestation. Other causal associations are not as clear but still very probable, such as the relationship between demographic factors and urban/road infrastructure with deforestation. From this point of view, population density and urban/road infrastructure can be seen as drivers of deforestation in the Ribeira Valley.

On the other hand, concerning the correlations between socioeconomic factors and deforestation it is more difficult to establish the causal relationships and, especially, its direction. In some cases, the deforestation could have generated income, therefore improving socioeconomic conditions of the population in the census tract.

[7] This qualitative model can be considered a form of path analysis model, in which the independent variables (drivers) have both direct and indirect effects on the dependent variable (deforestation).

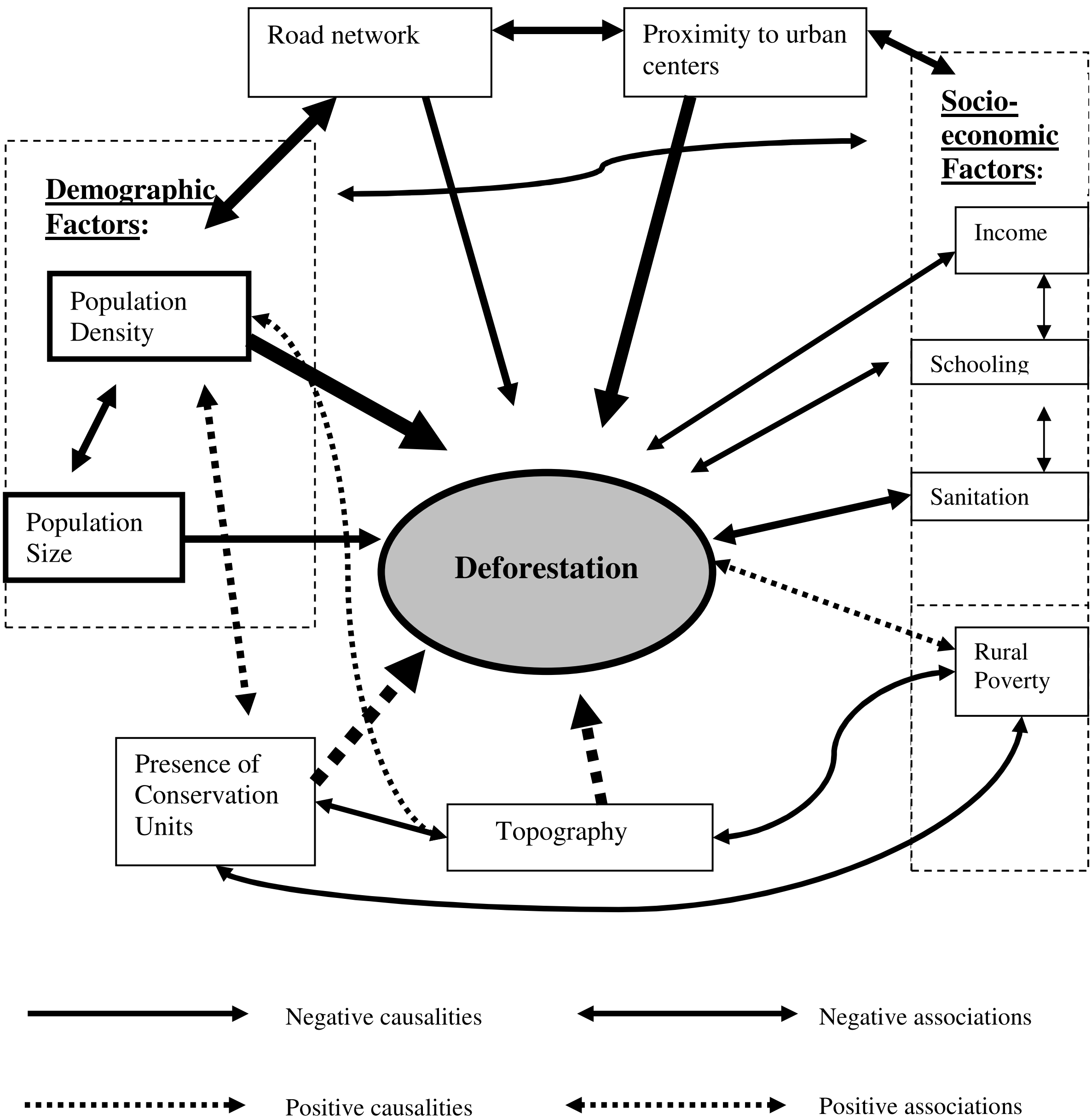

FIGURE 5. Qualitative model of correlation and causality between socio-demographic factors, topography, access to infra-structure, conservation units and deforestation in the Ribeira Valley (network of relationships among the independent variables and deforestation rates of rural census tracts).
(Sources: IBGE. Demographic Census 1991 and 2000 and Landsat TM images, 1990 and 1999 (scene 220/77)).

At this point, it is important to say that in our study on the Ribeira Valley, we first aimed to work with a regression model in order to identify and analyze the factors that could explain deforestation rates in the Ribeira Valley. However, we found out that the regression analysis was not appropriate to deal with all the independent variables selected because of the occurrence of strong correlations among them, what is known as multicollinearity. **Appendixes 1 and 2** show some regression models that we first tried to build for the Ribeira Valley[8].

Consequently, due to these restrictions to work with several independent variables in regression models, we decided to work, instead, with a qualitative model that could represent the network of relationships between the independent variables and the deforestation rates (the dependent variable). The major advantage of adopting a qualitative model is the possibility to map and represent graphically the diversity of factors associated with deforestation in the Ribeira Valley.

In future studies, we plan to build multivariate statistical models in order to perform quantitative analyses that are able to deal with all the selected independent variables of the qualitative model.

DISCUSSION OF RESULTS AND FINAL REMARKS

As we have seen, most of the deforestation models seen on the literature are empirical and one of the methodologies most used are regression analysis. However, as one can see in the discussion in the first section, the results found by deforestation regression models are problematical and contradictory. One of the most important limitations on deforestation regression models is the occurrence of strong correlations among the independent variables. Therefore, the deforestation regression models reviewed have strong limitations to address and untangle the role of demographic, socioeconomic and infrastructural variables to explain the deforestation processes that

[8] One can see in **Appendix 2** that the first regression model that we tried to build, with all the factors associated with deforestation (eleven independent variables), presented strong multicollinearity among the independent variables. In this first model, only two independent variables were significant – population density and presence or not of conservation units. In this first regression model, the R-squared was 0.458.
Based on the results for this first regression model, we built a second regression model (see **Appendix 3**) with only the two main factors associated with deforestation, namely population density (with positive association) and presence or not of conservation units (with negative association). Moreover, this last independent variable is a dummy, in which 0 is the absence and

they are studying. Because of these restrictions in regression models, we decided to work, instead, with a qualitative model that could represent the network of relationships between the independent variables and the deforestation rates in the Ribeira Valley.

In our qualitative model, the factors positively associated with recent deforestation in the Ribeira Valley are population size, density and growth, socio-economic conditions (levels of income, schooling and sanitation) and access to urban markets and road infra-structure. The major drivers of deforestation are population density and proximity to urban centers. Besides, these two factors, together with road density, are very correlated with each other, implying that population density might be a proxy to access to urban markets and road infra-structure.

On the other hand, the main factors negatively associated with deforestation are topography and the presence of conservation units. As mentioned before, the rural census tracts inside conservation units show much lower deforestation rates and significantly higher proportion of forest remnants than the census tracts outside the conservation units.

Therefore, it seems clear that since its implementation in the 1980´s, the conservation units - together with topography - have been the most important barrier to deforestation of Atlantic Forest remnants in the Ribeira Valley. However, the conservation units are distinguished by very low population densities and low levels of living conditions. Thus, if, on one hand, the conservation units have been quite successful in the preservation of forest remnants, on the other hand there has been a great out-migration from these areas and the perpetuation (or even worsening) of poverty and low socioeconomic conditions for the population still living inside or nearby the conservation units.

In this sense, some questions raised are: What type of environmental conservation program is taking place in the Ribeira Valley? Are poverty and out-migration of the local population, living in conservation units, a prerequisite for the preservation of Atlantic Forest remnants? Is this type of environmental conservation sustainable?

At last, it is important to mention that our analysis [of the relationships between deforestation rates and the independent variables] was not able to incorporate all the complexity of the factors involved in the processes of land cover change and

---

1 is the presence of conservation units. It is interesting to note that even with only these two

deforestation in course in the Ribeira Valley region. As discussed in the literature, deforestation processes do not proceed linearly. In other words, they are not dependent on one exclusive factor (e.g. population growth or road construction), nor are ahistorical. Instead, it is a combination of many factors (social, economic, demographic, political, institutional etc.), operating in different spatial and temporal scales and interacting in specific environmental, social and historical contexts.

For Lambin (1997: 389), "the most fundamental obstacle to progress in the understanding and prediction of human impacts on terrestrial ecosystems lies in the lack of a comprehensive theory of land-use changes. The role of a theory is to explain experimental findings and to predict new results". In this sense, there has been a great effort from the scientific community in the search for new theories and methods of analysis that could enable a better balance between geographic coverage, analytical precision and realism for the analysis and models of deforestation.

In our study on the Ribeira Valley, we also faced a dilemma to find the best balance between geographic coverage and analytical precision. On one hand, the simplicity of the correlation analyses and qualitative model limited the analytical power of our analysis of the factors associated with deforestation. But, on the other hand, these analyses enabled us to work with a relatively large numbers of variables and to incorporate all the rural census tracts of the Ribeira Valley (109 census tracts). If we used a regression analysis we would not be able to work with this large and diverse number of variables. Moreover, this geographic coverage would not be possible to achieve in a case study using fieldwork and in-depth analysis.

Therefore, our analysis of the factors associated with deforestation in the Ribeira Valley was able to incorporate three important aspects: 1) it presents a wide geographic coverage; 2) it uses very disaggregated spatial unit of analysis (census tracts); and 3) it integrates a large and diverse number of variables for the analysis (census variables, remote sensing/land cover variables and other spatial variables as topography and road infrastructure).

In this sense, maybe the most significant contribution of this study is the application of a methodology that integrates census and remote sensing variables, all information aggregated at the level of census tracts, for the development of an analysis of the associations between socio-demographic factors and deforestation. Hence, this is

---

independent variables the R-squared is 0.426.

one of the first studies in the field of “Population and Environment” to use this kind of methodology at the level of census tracts.

REFERENCES AND BIBLIOGRAPHY

AGUIAR, CÂMARA & ESCADA (2007). Spatial statistical analysis of land-use determinants in the Brazilian Amazonia: Exploring intra-regional heterogeneity. *Ecological Modelling, 209*(2-4), 169-188.

ALLEN J. C. & BARNES D. F. (1985). The causes of deforestation in developing countries. *Annals of the Association of American Geographers, 75*, 163–184.

ANGELSEN, A. & KAIMOWITZ, D. (1999). Rethinking the Causes of Deforestation: Lessons from Economic Models. *The World Bank Research Observer, 14*(1), 73-98.

BILSBORROW, R.E. & HOGAN, D. (eds). (1999). *Population and Deforestation in the Humid Tropics*. Liège: International Union for the Scientific Study of Population.

BRONDIZIO, E. S. (2000). Coping with External Forces: a nested-ethnographic approach to studying influences of conservation and development policy upon local communities in the Atlantic Forest of Brazil. *Proposal submitted to National Science Foundation*. July, 2000.

BRONDIZIO, E. S. *et al*. (2002). The Colonist Footprint: towards a conceptual framework of deforestation trajectories among small farmers in Frontier Amazônia. In WOOD, C. *et al* (eds.). *Patterns and Processes of Land Use and Forest Change in the Amazon*. Gainsville: University of Florida Press.

BRUNER, A. G. *et al*. (2001). Effectiveness of Parks in Protecting Tropical Biodiversity. *Science*, *291*(5), 125-128.

CAPISTRANO, A. D. & KIKER, C. F. (1995). Macro-scale economic influences on tropical forest depletion. *Ecological Economics 14*(1), .21-29.

CARR, D. L. (2004). Proximate Population Factors and Deforestation in Tropical Agricultural Frontiers. *Population and Environment*, *25*(6), 585-612.

CARR, D. L.; SUTER, L. & BARBIERI, A. (2005). Population dynamics and tropical deforestation: state of the debate and conceptual challenges. *Population and Environment*, *27*(1), 89-113.

CHOMITZ, K. M. & GRAY, D. A. (1996). Roads, land use, and deforestation: a spatial model applied to Belize. *World Bank Economic Review 10*, 487-512.

DAEE (1998). Bacia Hidrográfica do Ribeira do Iguape: Relatório Síntese do Plano de Ação. São Paulo.

DeMERS, M. N. (1997). *Fundamentals of geographic information systems*. John Wiley & Sons, Inc. New York.

EVANS, T. & MORAN, E. (2002). Spatial Integration of Social and Biophysical Factors Related to Land cover Change. In LUTZ, W. *et al*. Population and Environment: Methods of Analysis. *Population and Development Review*. A Supplement to Volume 28. Population Council, New York.

GEIST, H. & LAMBIN, E. F. (2001). *What drives tropical deforestation? A meta-analysis of proximate causes and underlying sources of deforestation based on subnational case study evidence*. LUCC Report Series No. 4.

GEIST, H. & LAMBIN, E. F. (2002). Proximate Causes and Underlying Driving Forces of Tropical Deforestation. *BioScience*, *52*(2), 143–150.

HOGAN, D.; CARMO, R.; ALVES, H & RODRIGUES, I. (1999). Sustentabilidade no Vale do Ribeira (SP): conservação ambiental e melhoria das condições de vida da população. *Ambiente e Sociedade*, Ano II, No 3 e 4, 151-175, Campinas. Núcleo de Estudos e Pesquisas Ambientais (NEPAM), Unicamp.

IGBP-IHDP Land Use Cover Change Project (LUCC). Available at <http://www.ihdp.uni-bonn.de/html/projects/i-lucc.html>.

Instituto Socioambiental (ISA). (1998). *Diagnóstico Socioambiental do Vale do Ribeira*: Documento Síntese. São Paulo.

JENSEN, J. R. (1996). *Introductory Digital Image Processing*. (2nd. Edition). New York: Prentice-Hall.

KAIMOWITZ, D. & ANGELSEN, A. (1998). *Economic models of tropical deforestation: a review*. Bogor, Indonesia: CIFOR.

LAMBIN, E. F. (1994). *Modelling Deforestation Process*. JRC-ESA, TREES Research Report # 1.

LAMBIN, E. F. (1997). Modeling and Monitoring Land-Cover Change Processes in Tropical Regions. *Progress in Physical Geography*, *21*(3), 375-393.

LAMBIN, E. F. *et al*. (1999). *Land-Use and Land-Cover Change: Implementation Strategy*. Stockholm, Sweden, IGBP Secretariat.

LAMBIN E. F. *et al*. (2001). The Causes of Land-Use and Land-Cover Change: Moving Beyond the Myths. *Global Environmental Change: Human and Policy Dimensions, 11*(4), 261–269.

LINO, L. F. (ed.). (1992). *Reserva da Biosfera da Mata Atlântica*. Plano de Ação. Vol. 1: Referências Básicas, junho/1992. Consórcio Mata Atlântica e Universidade Estadual de Campinas.

LIVERMAN, D. M. *et al*. (1998). *People and Pixels: Linking Remote Sensing and Social Science.* Washington, D.C: National Academy Press.

LUTZ, W. *et al*. (2002). Population and Environment: Methods of Analysis. *Population and Development Review*. A Supplement to Volume 28. Population Council, New York.

MATHER, A. S. & NEEDLE, C. L. (2000). The Relationships of Population and Forest Trends. *The Geographical Journal*, *166*, 2-13.

McCRACKEN, S. *et al*. (1999). Remote sensing and GIS at farm property level: Demography and deforestation in the Brazilian Amazon. *Photogrammetric Engineering and Remote Sensing, 65*(11), 1311-1320.

McCRACKEN, S. *et al*. (2002). Land-use Patterns on an Agricultural Frontier in Brazil: Insights and Examples from a Demographic Perspective. In WOOD, C. *et al* (eds.). *Patterns and Processes of Land Use and Forest Change in the Amazon*. Gainsville: University of Florida Press.

MERTENS, B. & LAMBIN, E. (1997). Spatial modeling of deforestation in Southern Cameroon: spatial disaggregation of diverse deforestation processes. *Applied Geography, 17*(2), 143-162.

OVERMARS, K.P., KONING, G.H. & VELDKAMP, A. (2003). Spatial autocorrelation in multi-scale land use models. *Ecological Modelling, 164*, 257–270

PEBLEY, A. R. (1998). Demography and the Environment. *Demography, 35*, 377-389.

PFAFF, A. S. P. (1999). What drives deforestation in the Brazilian Amazon? Evidence from satellite and socioeconomic data. *Journal of Environmental Economics and Management, 37*(1), 26-43.

ROCK, M. T. (1996). The stork, the plow, rural social structure and tropical deforestation in poor countries. *Ecological Economics, 18*(2), 113-131.

ROSERO-BIXBY, L. & PALLONI, A. (1998). Population and deforestation in Costa Rica. *Population and Environment, 20*(2), 149-185.

RUDEL, T. (1989). Population, development and tropical deforestation. A cross national study. *Rural Sociology, 54*, 327-338.

RUDEL, T. K. & ROPER, J. (1996). Regional Patterns and Historical Trends in Tropical Deforestation, 1976-1990: A qualitative comparative analysis. *Ambio, 25*(3), 160-166.

RUDEL, T. K. & ROPER, J. (1997). The Paths to Rain Forest Destruction: Crossnational Patterns of Tropical Deforestation, 1975-90. *World Development 25*(1), 53-65.

SCRIECIU, S. S. (2007). Can economic causes of tropical deforestation be identified at a global level? *Ecological Economics*, *62*(3-4), 603-612.

SOUTHGATE, D. *et al.* (1991). The Causes of Tropical Deforestation in Ecuador: A Statistical Analysis. *World Development 19*(9), 1145-1151.

WALKER, R. T. (1987). Land Use Transition and Deforestation in Developing Countries. *Geographical Analysis*, *19*(1), 18-30.

WALSH, S. J. *et al*. (1999). Scale-dependent relationships between population and environment in Northeastern Thailand. *Photogrammetric Engineering and Remote Sensing*, *65*(1), 97-105.

WOOD, C. H. & SKOLE, D. (1998). Linking satellite, census, and survey data to study deforestation in the Brazilian Amazon. In Liverman D. *et al*. (eds.) *People and Pixel: Linking remote sensing and social sciences*. Washington: National Research Council.

Appendix 1- Procedures for the classification of the satellite images and the generation of the land cover change variables

In order to produce land cover change variables, we used the following satellite images for the Ribeira Valley region:

- Landsat 5 TM (Thematic Mapper), orbit point 220/77 – October 23, 1990
- Landsat 5 TM (Thematic Mapper), orbit point 220/77 - September 21, 1999

We used the software ENVI 4.0 to process such images and initially registered the 1990 image, procedure later applied to the 1999 image. These images were geometrically corrected and registered with reference to the vector cartography of rivers and streams for the region (scale 1:10,000).

We used different digital image processing procedures to enhance the images regarding its vegetation, i.e., contrast enhancing, color composites, filters for special frequencies and mathematic operations for image classification. The color images that have presented the best results for visual interpretation were the ones produced with the TM4, TM5 and TM3 bands and channels red (R), green (G) and blue (B). We produced the color composites in 1990 and 1999 images, keeping the same contrast for both of them.

We performed a supervised classification based on training samples. Our basic source of information for the training samples were aerial photographs (scale 1: 8,000) of significant targets. Based on the largest sample possible (at least 5,000 pixels) in the two color composites (1990 and 1999), we started the process of supervised classification, using the classification algorithm known as MAXVER (maximum likelihood).

For each of the two images, we have distinguished seven land cover classes, which are: water, forest, mangrove, planted forest and non-forested areas (which include crops, pasture, bare soil and urbanized areas). However, for the purpose of this paper, we looked only at three main land cover types: water, non-forest and forest (which includes primary and secondary forest, mangrove and planted forest). Moreover, we have made no attempt to discriminate between primary and secondary forest, since it was not the paper's objective.

During the classification, we faced some problems with shadowing and illumination for the mountain part of the region. To solve these problems, we performed visual interpretation in some areas and corrected the classification with manual edition.

In order to test classification accuracy, we used a confusion matrix generating a Kappa coefficient of 0.9466 for 1990 image and 0.9442 for 1999 image. Such coefficient varies between 0 and 1, and the best classifications are those closest to 1. Therefore, the Kappa coefficients we have obtained are satisfactory, and we accepted both classifications.

After the classification was accomplished, we made a change detection analysis, in order to capture and quantify land cover changes between the two dates of the images (time period of 1990-1999). The most important land cover trajectories (or changes) we looked at were deforestation and preservation of forest remnants.

Finally, the two classified images were transferred to the software ArcGIS 8.1 and converted to a grid format in order to extract land cover classes, which were then aggregated to the spatial units of analysis, namely census tracts and municipalities.

Appendix 2: First Regression Model – considering all factors (independent variables) associated with deforestation

Dependent variable:

- PFLO_DESMA Deforestation rate between 1990 and 1999 (%)

Independent variables:

DENSID_200	Population Density in 2000
PBFCITY10K	Percentage of the census tracts area within 10Km from the nearest seat of municipality
PCBANHEIRO	Percentage of households with bathroom (in 2000)
MEDIA1_ANO	Average years of schooling of the heads of households (in 2000)
POPRESI	Population size of census tract in 2000
PBFROAD1KM	Percentage of census tracts area within the 1 Km buffers of road network
CRESC_91_0	Population Growth rate (1991-2000)
ELEV_RANGE	Variation of the elevation in the census tract
POBRES2000	Percentage of head of households in poverty (in 2000)
DUMMY_UC	Presence or not of conservation units (*dummy variable*)

REGRESSION

SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION

Dependent Variable	**PFLO_DESMA**	Number of Observations	109
Mean dependent var	14.9316	Number of Variables	11
S.D. dependent var	9.05654	Degrees of Freedom	98
R-squared	**0.458448**	F-statistic	8.29614
Adjusted R-squared	0.403188	Prob(F-statistic)	1.26494e-009
Sum squared residual	4841.63	Log likelihood	-361.419
Sigma-square	49.4044	Akaike info criterion	744.837
S.E. of regression	7.02882	Schwarz criterion	774.442
Sigma-square ML	44.4186		
S.E of regression ML	6.66473		

Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	16.68531	6.283559	2.655392	0.0092467
DENSID_200	**0.1796163**	**0.08622023**	**2.083227**	**0.0398338**
PBFCITY10K	-0.005465693	0.02571725	-0.2125302	0.8321371
PCBANHEIRO	0.01797709	0.03749503	0.4794527	0.6326831
POPRESI	0.0029043	0.001806728	1.607492	0.1111638
PBFROAD1KM	-0.03021472	0.0433779	-0.6965464	0.4877361
MEDIA1_ANO	0.05964127	1.156851	0.05155484	0.9589888
CRESC_91_0	0.00889898	0.0111959	0.7948428	0.4286249
ELEV_RANGE	-0.005654678	0.003978719	-1.421231	0.1584242
POBRES2000	-0.01954043	0.04344333	-0.4497912	0.6538528
DUMMY_UC	**-5.423307**	**2.016608**	**-2.689321**	**0.0084161**

REGRESSION DIAGNOSTICS

MULTICOLLINEARITY CONDITION NUMBER 28.48721

TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	25.38758	0.0000031

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	10	27.59633	0.0020942

Appendix 3: Second Regression Model – considering only two main factors (indep. variables) associated with deforestation – population density and conservation units

Dependent variable:

- PFLO_DESMA Deforestation rate between 1990 and 1999 (%)

Independent variables:

DENSID_200 Population Density in 2000
DUMMY_UC Presence or not of conservation units (*dummy variable*)

REGRESSION
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION

Dependent Variable	: **PFLO_DESMA**	Number of Observations:	109
Mean dependent var	: 14.9316	Number of Variables :	3
S.D. dependent var	: 9.05654	Degrees of Freedom :	106
R-squared	**: 0.426097**	F-statistic :	39.3501
Adjusted R-squared	: 0.415269	Prob(F-statistic) :	1.6536e-013
Sum squared residual:	5130.85	Log likelihood :	-364.581
Sigma-square	: 48.4043	Akaike info criterion :	735.162
S.E. of regression	: 6.95732	Schwarz criterion :	743.236
Sigma-square ML	: 47.072		
S.E of regression ML:	6.86091		

Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	15.60513	1.341112	11.63596	0.0000000
DENSID_200	0.2299799	0.06010256	3.826458	0.0002201
DUMMY_UC	-7.722522	1.570706	-4.916592	0.0000032

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER 3.76745
TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	43.38495	0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	2	5.499227	0.0639526
Koenker-Bassett test	2	2.696993	0.2596303